The Writing System

Introduction

The Persian alphabet, like the alphabets of many Muslim nations, is based on the 28-letter alphabet of Arabic. It has 32 letters and is written from right to left. Four of these letters are devised exclusively for Persian. We shall see these letters later. The Persian alphabet does not use capital letters. It does, however, use a modified version of the Western punctuation system.

Unlike the alphabet, the Persian numbers are written from left to right. Persian uses the same figures as Arabic, the language from which the Arabic numerals of English are derived. The Persian numbers, thus, can be easily compared with those in English. We shall discuss the numerals later.

In the pages that follow, the alphabet is divided into a number of letter groups. Each letter group uses a basic form. After explaining each basic form, and the manner in which that form is made, individual letters are introduced and discussed. These individual letters are different from the basic form in the number of dots and the type of diacritics or symbols that may accompany them.

The sequence used in the earlier stages of this study is not the same as the native sequence of letters taught in the schools in Iran. Once the students are familiar with the mechanism for producing letters and letter blocks, the native sequence will be introduced and memorized. This latter sequence is the one used in all alphabetizing of telephone directories, dictionaries, and other such manuals using the Persian alphabet.

The entire Persian alphabet is presented on the next page. The letters are in their independent or isolated form. The sequence is the native sequence mentioned above in relation to the use of dictionaries, directories, etc. The following information is provided for each letter: the Persian name of the letter, the pronunciation or sound that the letter represents,[1] the transliteration used to symbolize that letter in the Latin script (transliteration is an aid for the graduate student undertaking research using secondary materials in the field),[2] and information as to whether a given letter is a connector or a nonconnector. The symbol (+) marks a connector. The symbol (-) marks a nonconnecting letter. More information on connectors and nonconnectors will follow.

In the pages that follow, this list, and the characteristics of each letter or group of letters, will be discussed in detail. Before beginning the discussion of letter groups, however, there are three points that must be mentioned. First, not all the letters of the Persian alphabet connect to the letters that follow them. There are seven letters known as the nonconnectors. They connect only to connecting letters that precede them. Second, depending on where in a block of letters a connecting letter is used (it is necessary to make the distinction between a block of letters and a word, since most Persian words are made up of two or three blocks of letters where each block, except possibly the last, ends in a nonconnecting letter), the shape of the letter may undergo a substantial amount of reduction--in most cases a modified form of the initial portion of the letter is used. Third, a group of letters may share the same basic form. In such cases dots and other diacritics distinguish one letter from another.

The major distinction to be made is the connector/nonconnector distinction. Since the English alphabet does not make this distinction, the principle underlying letter blocks as formative components of words sometimes escapes American students.

Letter	Persian name	Pronunciation/ transcription	transliteration	connected/ nonconnected
ا	ælef	see vowel letters	a	-
ب	be	b	b	+
پ	pe	p	p	+
ت	te	t	t	+
ث	se	s	ṡ	+
ج	jim	j	j	+
چ	če	č	č	+
ح	he	h	ḥ	+
خ	xe	x	x	+
د	dal	d	d	-
ذ	zal	z	ż	-
ر	re	r	r	-
ز	ze	z	z	-
ژ	že	ž	ž	-
س	sin	s	s	+
ش	šin	š	š	+
ص	sad	s	ṣ	+
ض	zad	z	ẓ	+
ط	ta	t	ṭ	+

ظ	za	z	ẕ	+
ع	'eyn	'	' or ᶜ	+
غ	qeyn	q	ġ	+
ف	fe	f	f	+
ق	qaf	q	q	+
ك	kaf	k	k	+
گ	gaf	g	g	+
ل	lam	l	l	+
م	mim	m	m	+
ن	nun	n	n	+
و	vav	v / u	v / u	-
ه	he	h	h	+
ی	ye	y	y	+

The Nonconnectors

The first group of letters to be discussed is the nonconnectors. Obviously, the designation nonconnector is somewhat misleading. These letters actually connect only to connecting letters that precede them, never to letters that follow them. If a word is composed only of nonconnecting letters, there will be no block of connected letters. All the letters in such a word are independent, pretty much like printed English letters. It is only when the connecting letters enter the picture, and when they precede the nonconnectors that the latter use a hook (to the right) to attach to preceding connecting letters.

In the discussion that follows we shall deal with the independent form of the nonconnectors only. Later on, when we learn a few connecting letters as well, we shall return to the nonconnectors. The words that we produce at this stage do not include letter blocks, they are words consisting of independent nonconnecting letters only. Now, let us look at the seven nonconnecting Persian letters. The first line shows the nonconnectors in their isolated forms. The second line places them in relation to an imaginary line on the paper:

ا د ذ ر ز ژ و

ـا ـد ـذ ـر ـز ـژ ـو

Basic Forms

1. The Letter **ælef**

The letter ا **ælef** is a single vertical stroke. The independent form of **ælef** is written from top to bottom. It rests on the line (see below, for the final form of **ælef**).

basic form

ا

As previously noted, several letters of the alphabet may share the same basic form. Dots and other symbols are used to distinguish one member of such a group from another. Consider the next two sets of letters: **dal/zal** and **re/ze/že**. The former includes two letters differentiated by no dot and one dot, the latter includes three letters differentiated by no dot, one dot, and three dots respectively.

2. The **dal-**group

basic form

د

د **dal** always represents the sound **d**.

ذ **zal** is one of the four letters representing the sound **z**. **zal** is used primarily in words of Arabic origin and is used much less than the letter **ze**. ذ is transliterated as ż.

3. The **re-**group

basic form

ر

ر **re** always represents the sound **r**.

ز **ze** is one of the four letters that represent the sound **z**.

ژ **že** always represents the sound **ž**. It is found in words of non-Arabic origin. It is a Persian letter.

4. The letter **vav**

basic form

و

و **vav** has two basic uses. It is used as the vowel **u** as well as the consonant **v**. As the vowel **u**, when used syllable initially, it must be preceded by an **ælef:** او **u** (he). In syllable medial and syllable final positions, it usually represents the sound **u**, except if it is preceded by an **ælef**, in which case it is pronounced **v**: واو **vav**. Syllable initially (without a preceding **ælef**), it represents the consonant **v**. See also further below for the use of **vav** as a consonant.

In a few instances such as in دو **do** (two), **vav** represents the vowel **o**. The words that use **vav** to represent the vowel **o** will have to be memorized.

Vocalization

The letter **ælef** is most frequently used to represent the unwritten vowels of the Persian alphabet. In fact, in syllable initial position, with the help of a set of diacritics, the letter ا **ælef** represents almost all the vowel initial syllables of Persian. These symbols are:

1. Madda (◌̃). When added directly over the top of a syllable initial **ælef**, the madda represents a syllable initial **a**, thus: آ = **a**. Dictionaries explain the **madda** as a modified form of ا **ælef** itself. They say it is written in this way to avoid writing two **ælefs** (اا) to represent the initial vowel **a**. If the syllable is not a word-initial syllable, the **madda** may be dropped. The letter **ælef** alone represents the sound **a** in syllable-medial and syllable-final positions:

آر آز را آرا آزاد دارا داد

2. Kasra, or **zir** (◌ِ). When added directly underneath a syllable initial **ælef**, **kasra** represents a syllable initial **e**, thus: اِ = **e**. This symbol may be added directly underneath a consonant letter as well. Its value remains the same: دِ is pronounced **de**. The **kasra** is not usually used in nontextbook materials such as newspapers, etc. :

رادْ رازْ درازْ دژْ ادْرار

3. Fatha, or **zebar** (◌َ). When added directly over a syllable initial ا **ælef**, the **fatha** represents a syllable initial **æ**, thus: اَ = **æ**. This symbol may be added directly over a consonant letter as well. Its value remains the same: دَ is pronounced **dæ**. The **fatha** is usually not used in nontextbook materials such as newspapers, etc. Compare:

آز دَر رَد

4. Zamma, or **piš** (◌ُ). When added directly over a syllable initial ا **ælef**, the **zamma** represents a syllable initial **o,** thus: اُ = **o**. This symbol may be added directly over a

consonant letter as well. Its value stays the same: دُ is pronounced **do**. The **zamma** is usually not used in nontextbook materials:

اُردو رُز دُر

We have already seen the combination of **ælef** and **vav** representing an initial **u**: او . We have also mentioned that in most syllable medial and syllable final positions the letter **vav** alone symbolizes the sound **u** :

او دور اُردو

The only remaining vowel to be discussed is the vowel **i**. Since the case of **i,** like that of **u,** is one of letter combinations (in this case **ælef** + **ye** in syllable initial, and **ye** alone in syllable medial and syllable final positions), we shall postpone the discussion of this vowel until later (see Letters with the Arabesque below the Line).

5. Sokun. In English, it is the presence rather than the absence of a vowel that is important. In the case of vowels of Persian, however, the absence of a vowel is also significant. In fact there is a symbol (◌ْ) called **sokun** to represent the absence of a vowel in syllable medial and syllable final positions. The last letter of the word is usually not marked for **sokun**:

اُرْدو دُزْد دَرْد آرْد

Before concluding this section on the independent forms of the nonconnectors and the discussion of vocalization, it should be added that the rules mentioned above are not a hundred percent workable at all times. We mentioned, for instance, that although the **zamma** represents the sound **o,** in a few cases this sound is produced with a **vav** as well: دو **do**. Similarly, while the use of the symbols mentioned above is generally limited to textbooks, in a few cases these symbols play an essential role in differentiating words which otherwise would remain undistinguishable. Compare:

	دَرْد	**dærd**	pain
as opposed to			
	دُرْد	**dord**	dregs

We shall, in the course of our explanations, return to these inconsistencies and point them out. Compare:

آوَرْد واو اَژْدَر اُرْدو

اِدْرار دُزْد رُز

Study the structure of the following words carefully. Note that all the letters are nonconnectors. There are, therefore, no connected letter blocks involved. All the words are composed of independent letters. The dots in the transcriptions (**.**) represent syllable boundary:

روز **ruz** (day). The letter و in syllable medial position represents the vowel **u**.

دارا **da.ra** (title of ancient Iranian kings). In syllable final position, the vowel **a** is written with an **ælef**.

او **u** (he/she). Syllable initially, the sound **u** is produced with the letter **vav** preceded by an **ælef**.

آرْد **ard** (flour). Syllable initially, the vowel **a,** is composed of **ælef** with a **madda** on top of it. A **sokun** on the letter **re** shows that this letter is not voweled.

وُرود **vo.rud** (entrance). Syllable initially, the consonant **v** is produced with a و alone. The vowel **o** is written with a **zamma** added above the preceding consonant. The vowel **u** is produced with a و in syllable medial position.

دو **do** (two). One of the words in which و represents **o**

دَر **dær** (door). The vowel **æ** is produced with a **fatha** over **dal**.

دژ **dež** (fort). The vowel **e** is written with a **kasra** underneath **dal**.

دُرْد **dord** (dregs). **o** is written with a **zamma**. **sokun** over **re**.

Homework

1. Copy the following and hand in to your instructor:

زور رود دارو اَژْدَر آرِزو

دَرْد دارَد دُرود

دراز روژ زر راز اُردو

داد راد زاد ذو

2. Write the following in the Persian script and hand in to your instructor:

az, ra.dar, zærd, dozd, važ, do, vav, da.dar, żu
ru, va.dar, dud, dord, ed.rar, dar, dær, da.ræd

The Connectors

The connectors are letters which can join a preceding letter, a following letter or both. These letters will be discussed in four groups as follows:

1. Letters that in full form are always made above the line.
2. Letters, in full form, with an arabesque made below the line,
3. Letters with a reverse arabesque, and
4. The letter mim.

Like the nonconnectors, the connectors have an independent form. This is the full form of each letter. Depending on where in a block of letters a certain letter is used, the shape of the independent letter may be reduced or somewhat modified. These modifications do not affect the number of dots or other diacritical symbols attached to the independent form of the letter. They affect the basic form only.

In order to study the letters that form a given word, the word is broken down into its formative letter-block components. For each letter block the following letter shapes, or allographs, are distinguished:

a. initial, the form of the letter that begins a letter block and connects to the following letters only;
b. medial, the form of the letter that joins both the preceding and the following letters;
c. final, the form of a connecting or nonconnecting letter which joins the preceding letter only;
d. independent, the form of the letter which follows nonconnecting letters when only one letter is remaining to be written.

As an example of the use of allographs in the positions mentioned above, let us consider the letter group usually referred to as the **be**-group. This group consists of four letters distinguished by dots only; it is a member of the larger group of letters that are always written above the line.

1. Letters that in full form are made above the line

a. The **be-** group

	basic form
initial	
medial	
final	
independent	

ب **be** always represents the sound **b**

پ **pe** always represents the sound **p**. One of the four Persian letters, it is found in words of non-Arabic origin.

ت **te** is one of two letters representing the sound **t** (see the letter **ta**)

ث **se** is one of three letters representing the sound **s**. It is found in words of Arabic origin. ث is transliterated as **ṡ**.

Note: Normally each one of the sounds of a language is represented by one letter, or by a combination of letters, in the orthographic system of that language. In the case of the Persian orthography, however, due to the adaptation of the Arabic script, some sounds such as **t** and **s** are represented by more than one symbol. In fact, a minor example of this can be seen in English where two letters (**k** and **c**) represent the sound **k**. For our purpose, and at this early stage, we can ignore the reasons why such a multiplicity of letters should represent a single sound. The introduction into the language of Arabic words with their "frozen" orthographic form could explain part of the problem.

The knowledge of which letter to use is part of a number of things that one learns about a word. These include the pronunciation of the word, its meaning, and whether it is a borrowed form. The latter is a major factor in guiding the student in his choice of the correct spelling for a given word; it is, however, not the only one. For the present, however, the transliteration system outlined above will aid us in distinguishing these letter-sound combinations.

We said earlier that the designation "nonconnector" was misleading, and that the non-connectors actually join connecting letters that precede them to their right. Now that we have learned a few connecting letters as well, let us see how the system works.

In order for a nonconnecting letter to join a connecting letter, a connecting "hook" is attached to the right side of the independent form. This is, in fact, the same process that converted the initial **be-** into medial, and the independent **be** into final. For **ælef** this hook is at the bottom, to the right. It rests on the line. A similar hook joins the other members of the

group to the letters that precede them as they come into contact on the line. In the following, the arrows show the direction of the movement of the pen:



Here are some examples of connecting and nonconnecting letters forming letter blocks and words:

bu	بو	ب + ـو
ta	تا	ت + ـا
pa	پا	پ + ـا
tæb	تَب	تَ + ـب
tæbær	تَبَر	تَ + ـبَ + ـر
torbæt	تُرْبَت	تُ + ـرْ + بَ + ـت
torab	تُراب	تُ + ـر + ا + ب
tur	تور	ت + ـو + ر
pedær	پِدَر	پِ + ـدَ + ر

b. The Letter **fe**

The independent form of this letter is similar to that of the letter **be** just discussed. The difference lies in a loop that appears on the initial portion of the letter **fe**. The dot for **fe** is placed right above this loop. The body of the letter rests on the line. Here are the basic forms of **fe**:

	basic form
initial	فـ
medial	ـفـ
final	ـف
independent	ف

ف **fe** always represents the sound **f**.

c. The **kaf**-group

This group, too, resembles the **be-**group; however, instead of a loop added onto the initial portion of **be,** the initial portion is elongated upwards. The **kaf-**group does not use any dots, but it uses strokes.

The two members of this group are called **kaf** and **gaf**. **kaf** is always written with one stroke. This stroke is optional on the independent **kaf** and may not appear in printed materials. When writing, the stroke is drawn from top to bottom left where it meets the top of the initial portion of the letter. **gaf** has two obligatory strokes. They are made in the same way that the stroke for **kaf** is made. They look like two small horizontal lines with the top line somewhat smaller. Inside the independent forms of both of these letters, there is a "squiggle," which is found in the printed materials quite often. This squiggle does not have any specific meaning. It is purely decorative and may be left out altogether. Like **be** and **fe,** the bodies of the letters **kaf** and **gaf** rest on the line. Here are the positional variants of the basic form for these letters:

	basic form
initial	كـ
medial	ـكـ
final	ـك
independent	ك

ك **kaf** (with one stroke) always represents the sound **k**. In print, on the final and independent forms the stroke is optional.

گ **gaf** (with two obligatory strokes) always represents the sound **g**. This letter is found in words of non-Arabic origin. **gaf** is a Persian letter.

d. The **ta**-group

Were the initial portion of the basic form for **be** to meet the end portion of that letter, a loop would result. The addition of the stroke of **kaf** vertically to this loop would result in the basic form for **ta** and **za**. The two members of the **ta**-group are distinguished by one dot. The base of the loop rests on the line.

	basic form
initial	
medial	
final	
independent	

ط **ta** is one of two letters representing the sound **t**. When transliterating this letter, a dot is placed underneath **t** : **ṭ**

ظ **za** is one of four letters representing the sound **z**. When transliterating this letter, a line is placed underneath it: **z̲**. Both of these letters are found in words of Arabic origin.

e. The letter **he**

The letter **he** does not have any dots. Deriving this letter from the basic form of **be** would involve some stretching of the imagination. That is, if we have not already gone too far with **ta**. Here are the positional variants of **he** :

	basic form
initial	
medial	
final	
independent	

ه **he** is one of two letters representing the sound **h**. This letter is usually referred to as

he-ye hævvæz (see the section on "letters and numbers"), or **he-ye hendune**. We shall see later that the word **hendune** means water-melon, and that ه is the letter with which this word is written in Persian.

Study the structure of the following words carefully:

تَگَرْگ **tægærg** (hailstone). In this word we have an initial تـَ (**te** with a **fatha** for the vowel **æ**). This is followed by a medial ـگَـ (**gaf** marked with another **fatha**), again representing the vowel **æ**. The block is completed by a final ـرْ **re**. Since **gaf,** a connector, is the only letter to be added to a nonconnector, it is used in its independent form. Lack of vowel at the transition of **re** to **gaf** is marked by a **sokun**.

بَرْف **bærf** (snow). In this word we have an initial بـَ (**be** marked with a **fatha** for the vowel **æ**). The short block ends in a final ـرْ **re** marked with a **sokun**. ر **re,** a nonconnector, is followed by an independent ف **fe,** because this letter is the last and only letter to be added after a nonconnector. Once again, **sokun** marks the lack of a transition vowel between **re** and **fe**.

پُتْك **potk** (sledge hammer). This is a one-block word with an initial, medial and final form. We have initial پـُ (**pe** marked with a **zamma** to represent the vowel **o**), a medial ـتْـ (**te** marked by **sokun** for no vowel), and a final ـك **kaf**.

رَهْبَر **ræh.bær** (leader). Independent رَ (**re** [nonconnecting] marked with a **fatha**), block initial هْـ (**he** marked with a **sokun**), medial ـبَـ (**be** marked with a **fatha**), and a final ـر **re**.

بِهْبود **beh.bud** (well-being). Initial بِـ (**be** marked by a **kasra** for the vowel **e**). This is followed by a medial ـهْـ (**he** marked with a **sokun**). The block continues with a medial ـبـ **be** followed by a final ـو **vav** indicating the vowel **u**. The letter **vav** ends this block. There is only one more letter left to be written, a د **dal**. The **dal** is written in its independent form.

Homework

1. Copy the following and hand in to your instructor. Compare your pronunciation with section (3) below:

کَدِر	پوك	توپ	فوت
گُرْگ	بَرْگ	ظَفَر	وَفا
گور	طَرَف	تور	فَرْدا
کَف	تَه	دُکْتُر	تَب
بِهْتَر	تاب	رَگ	اِرْتِباط
کورِه	فَتْوا	بِهْدار	رُتْبِه

2. Write the following in the Persian script and hand in to your instructor. The sounds symbolized by more than one letter are transliterated. Compare your answers with section (4) below:

bot	ku.zeh	fækk	tut
gærd	gord	gerd	beh
ru.deh	du.deh	bæhr	a.zad
ba.zu	žærf	guž	dæf
zeh	bu	bæd	ah
ab	æž.dær	va.zheh	a.hæk
a.har	ahu	or.du.gah	ræbṭ
kæbk	ut	śær.væt	erś
żat	keżb	kæś.ræt	

3. Write the following in the Persian script. Compare your renditions with section (1) above.

kæ.der	puk	tup	fut
gorg	bærg	ẓæ.fær	væ.fa
gur	ṭæ.ræf	tur	fær.da
kæf	tæh	dok.tor	tæb
beh.tær	tab	ræg	er.te.baṭ
ku.reh	fæt.va	beh.dar	rot.beh

4. Read the following aloud. Compare your pronunciation with section (2) above.

بُت کوزِه فَكّ توت

گَرْد گُرْد گِرْد بِه

رودِه دودِه بَهْر آزاد

بازو ژَرْف گوژ دَف

زِه بو بَد آه

آب اَژْدَر واژِه آهَك

آهار آهو اُرْدوگاه رَبْط

كَبْك اوت ثَرْوَت اِرْث

ذات كِذْب كَشْرَت

فانی خَسْتِگی تِشْنِگی مِدادی اِمْدادی

اُسْتادی شاگِرْدی بیکاری بیماری

چای بوی موی سوی

c. independent or final **vav** for **u**. This letter represents the sound **v** in syllable initial, or in syllable final position after **a**. There are a few exceptions such as لَـغْـو **læqv** and سَـــرْو **særv**. These we shall discuss later.

Please read aloud:

خُروج دور کِدو عَدو رود دوش زور

روز روژ روزِه کَبود خوشْبو گو مو

پوك پول کود سود سوژِه عوعو

سوسو لولو کوکو راوی ناو گاو واك

واو وام وان وای ویران وِیْس وَنَك

واسِطه

3. Double consonants are marked with a **tæšdid** (ّ) placed directly over the consonant geminated, or doubled. Please read and identify words with doubled consonants:

بُت بَد بَر بُز بَعبَع حَبّ بَس بُن جا

جَدّ جُز جَم چَپ دارو وُرود دادار رو

دو حَجّ خَر خَز خُل خَم خُم حَدّ خو

دِقّ دِل حَلّ حَكّ حِسّ حَیّ دَم دُم دَه

سَدّ سَر سَگ سَمّ سو شَكّ شَطّ سی

دِه صَد شَل شُل طِیّ عَرْعَر صَف صُم

غُرْغُر غَم فَنّ گَر كُر كُلّ لَعلَع وَغوَغ

هُد هُد سیر سِرّ كَرِه كُرّه مَزّه

4. Ligatures are combinations of two letters written in a "frozen," stylized form. In Persian, **lam** and **ælef** represent such frozen forms. The independent ligature looks like this: ﻻ , and the final looks like this: ﻼ . Note that even though **lam** is a connecting letter and technically should be connected to the following **ælef**, the **ælef** is written as an independent stroke, slightly slanted and placed inside the **lam**. Because the ligature ﻼ ends in an **ælef**, the letter that follows does not join it.

When an **ælef** follows a **kaf** or a **gaf**, this combination results in a ligature as well. In this case the beginning of the **kaf** or **gaf** is written (actually drawn) slightly curved towards the **ælef**, thus:.

Please read aloud:

قَد كَج كَف گُلا بی كَم كِه كی كِی

كِلاس یِیْلاق گَپ گَچ گَز گُل گِل

کُلاه لار لاس لُر لَب لَج ما مُچ مُخ

لاف مَد مَدّ مو مَه مِه یار لاك فِلات

مَلال مِی نَخ اِصْطِلاحات مَجَلاّت مَلامَت

هَم کاج کاخ عام ساق ناف نان هاری

یاس ایالات ایلات

5. The letter **vav** again:

After **xe**, the vowel **a** is sometimes written as a silent **vav** followed by an **ælef**: خوا **xa**. Thus, **xastæn** "want" is written خواسْتَن ; **xab** "sleep" is written خواب . This results in homonyms with different spellings. Example:

خوار	**xar**	contemptible
خار	**xar**	thorn

and in identical spellings with different pronunciations:

خوار	**xar**	contemptible
خَـوارِج	**xævarej**	foreigners

The second form may be distinguished by placing a **fatha** over **xe**: خَوارِج . The combination of **vav** and **ælef** discussed above can represent **a** only after **xe**. Therefore, عَوامِل should be read **'æ.va.mel**, never **'amel**.

Please read the following:

ایمان ویران خوان کور خوانْدَن

خواهِش

بَد خو خان عَدو خواسْت خواهَر اوت

6. The Diphthongs **ey** and **ow**:

The diphthongs **ey** and **ow** are indicated by letters **ye** and **vav** with a **fatha**. Thus:

سَیف	**seyf**	sword
بِین	**beyn**	between
دَور	**dowr**	around

The use of the **fatha** is primarily to distinguished **ey** from **i**, and **ow** from **u**. Compare:

بِین	**beyn**	between	دَور	**dowr**	around
دور	**dur**	far	بین	**bin**	pres. stem of "to see"

The spelling of **ey** and **ow** are clearly not very accurate phonetically. This is because most words containing **ey** and **ow** were borrowed from Arabic. During the borrowing, the pronunciation changed to fit the Persian system, but the orthography remained unchanged. Occasionally these words are pronounced as they are pronounced in Arabic (e.g., **sayf** or **lawn**) to demonstrate the speaker's learning.

7. The **Tanvin**

The Arabic indefinite accusative suffix has retained its original form. Thus a few Persian words end in an **ælef** on which a **tænvin** mark (ً) is placed, thus: اً The **tænvin**, the seat of which is usually an **ælef**, is pronounced **æn**. Example:

فوراً	fow.ræn	quickly	مَعْمولاً	mæ'.mu.læn	usually
				(**lam** + **ælef**, independent)	
تَـقـْریباً	**tæq.ri.bæn**	approximately	فعْلاً	**fe'.læn**	now
				(**lam** + **ælef**, attached)	
جِدّاً	**jed.dæn**	seriously	اَبَـداً	**æ.bæ.dæn**	at all

Certain Western words like تلگرافاً **tel.ge.ra.fæn** "by telegram," are also written with a **tænvin**.

8. The Short **ælef**

Certain Arabic words with final **ye** have retained this feature after their adoption into Persian. Thus, موسی is read **mu.sa** instead of **mu.si**, and حتّـی is read **hæt.ta**.

9. The Letter **hæmza**

The **hæmza** represents the glottal stop. In writing, it usually appears in conjunction with an **ælef**, a **vav** or this special seat : ـئـ (ye without dots). Examples:

مَأخَذ	**mæ'.xæz**	source
مَسْئول	**mæs.'ul**	responsible
رِئیس	**re.'is**	director
هیئت	**hey.'æt**	delegation
مُؤمِن	**mo'.men**	a pious person
مؤبَد	**mo'.bæd**	Zoroastrian priest

Please read the following:

رَئوف تِئوری تَأکید اِمْریکائی مَأمور

اِسْپانیائی اِمْضاء اِمْلاء قاآن مَآب

اَلآن سوء قُرْآن جُزْء جاهائیکه تَأَلُّم

مَعْمولاً فِعْلاً اَصْلاً مَثَلاً

10. Morphemes that may join other forms

Some prefixes and some monosyllabic words tend to join the word that follows them in writing. Among the most frequently used forms that join the words following them are **mi-** and **be-**. **-ra** joins the word that precedes it (see also contractions, below). The number of forms that can be used as independent forms or joined to following forms is indefinite. The student, after working with the language for a while , will have a better grasp of the situation. For this reason here we shall give only a few examples:

میکُنَم = می کُنَم	**mi.kon.æm**	I do
اینرا = این را	**in.ra**	this (object)

11. Contractions

When two words are combined, sometimes one of the original letters of a word may be dropped:

مَن را = مَرا	**mæn ra = mæra**	I (object), me
تو را = تُرا	**to ra = tora**	you (sing., object)
بِه هَم = بِهَم	**beh hæm = behæm**	to each other
این اَسْت = اینَسْت	**in æst = inæst**	this is

Please read the following and identify:

1) the morphemes that have joined other words
2) the contractions, if there are any:

میزَنَم بِاو می‌دِهَم یِکْد یگَر بِیکْد یگَر

بِهَمْد یگَر کِتابْرا بِخانِه بِا صْفَهان مَنْرا

مَرا تُرا اینَسْت

Numerals

The Persian numbers are written from left to right. They resemble the Arabic numbers which originate in the Hindu numerical system. The numbers from zero to ten are presented below.

It was mentioned earlier that the Persian numbers resemble those of English. An attempt is made here to show the relationship of the two sets of numbers by deriving the Persian set from English. After this the student should be able to remember the shape of the Persian letters with ease. These are the cardinal numbers:

printed	0	1	2	3	4	5	6	7	8	9
written										
stage 1										
stage 2										
written										
printed	۰	۱	۲	۳	٤	٥	٦	٧	٨	٩

The Cardinal Numbers

English	**transcription**	**written**	**figures**
zero	sefr	صفر	۰
one	yek	يك	۱
two	do	دو	۲
three	se	سه	۳
four	čæhar	چهار	٤
five	pænj	پنج	٥
six	šiš	شش	٦
seven	hæft	هفت	٧
eight	hæšt	هشت	٨
nine	noh	نه	٩
ten	dæh	ده	۱۰
eleven	yazdæh	يازده	۱۱
twelve	dævazdæh	دوازده	۱۲
thirteen	sizdæh	سيزده	۱۳
fourteen	čæhardæh	چهارده	۱٤
fifteen	panzdæh	پانزده	۱٥
sixteen	šanzdæh	شانزده	۱٦

seventeen	hivdæh	هفده	۱۷
eighteen	hijdæh	هیجده	۱۸
nineteen	nuzdæh	نوزده	۱۹
twenty	bist	بیست	۲۰
twenty-one	bisto yek	بیست و یك	۲۱
thirty	si	سی	۳۰
forty	čehel	چهل	٤۰
fifty	pænjah	پنجاه	٥۰
sixty	šæst	شصت	٦۰
seventy	hæftad	هفتاد	۷۰
eighty	hæštad	هشتاد	۸۰
ninety	nævæd	نود	۹۰
one hundred	sæd	صد	۱۰۰
two hundred	devist	دویست	۲۰۰
three hundred	sisæd	سیصد	۳۰۰
four hundred	čæharsæd	چهارصد	٤۰۰
five hundred	pansæd	پانصد	٥۰۰
six hundred	šešsæd	ششصد	٦۰۰
seven hundred	hæftsæd	هفتصد	۷۰۰
eight hundred	hæštsæd	هشتصد	۸۰۰
nine hundred	nohsæd	نهصد	۹۰۰
one thousand	(yek) hezar	هزار	۱,۰۰۰
one million	(yek) melyun	ملیون	۱,۰۰۰,۰۰۰

The Ordinal Numbers

The ordinal numbers are derived from the cardinal numbers with the addition of the suffix **-om**. This suffix is added to the last digit of the number:

یکم	yekom	first [3]
دوّم	dovvom	second
سوّم	sevvom	third
چهارم	čæharom	fourth
پنجم	pænjom	fifth
بیست و سوّم	bisto sevvom	twenty third
صد و نود و هفتم	sædo nævædo hæftom	one hundred and ninty seventh

Fractions

Fractions are formed by combining the cardinal numbers with the ordinal numbers. Example:

سه چهارم	se čæharom	three quarters (three fourths)
دو سوّم	do sevvom	two thirds

Percentage

Percentage is expressed with the word صد **sæd** (hundred) and a cardinal number. Example:

صدی بیست	sædi bist	twenty per cent.

Letters and Numbers

The numerical system known as the **æbjæd** is an aid for remembering historical dates pertaining to events of importance. The system is based on the Arabic alphabet and uses Arabic letters only. The system is used in Persian with the addition of four letters. It should be noted that the values given to the Persian letters **pe**, **če**, **že**, and **gaf** are the same as those already assigned to **be**, **jim**, **ze**, and **kaf** respectively. Here are the eight nonsensical words on which the **æbjæd** system draws:

هَوّز حُطّی کَلَمَن سَعْفَص قَرِشَت

اَبْجَد

ثَخّذ ضَظّغ

Each of the letters in the words listed above is assigned a number. The order of the words and, thereby, the value of each letter within the series remains a constant. The value assigned to each letter is provided below:

1	١	ا
2	٢	ب (پ)
3	٣	ج (چ)
4	٤	د
5	٥	ه
6	٦	و
7	٧	ز (ژ)
8	٨	ح
9	٩	ط
10	١٠	ی
20	٢٠	ك (گ)
30	٣٠	ل
40	٤٠	م
50	٥٠	ن
60	٦٠	س
70	٧٠	ع

80	۸۰	ف
90	۹۰	ص
100	۱۰۰	ق
200	۲۰۰	ر
300	۳۰۰	ش
400	٤۰۰	ت
500	٥۰۰	ث
600	٦۰۰	خ
700	۷۰۰	ذ
800	۸۰۰	ض
900	۹۰۰	ظ
1,000	۱۰۰۰	غ

In order to assign an **æbjæd** to a given date, the date is analyzed into various letter combinations until a phrase somewhat defining the event to be commemorated is achieved. The chronogram thus arrived at is then quoted to commemorate that event. For example, the chronogram for Nadir Shah's proclamation of the Persian throne is the Arabic phrase:

الخیر فی ما وقع **al-xayru fi ma væqæ'** (the best is in what happened). The total which corresponds to the year 1148 A.H. (1735-6 A.D.) is arrived at as follows:

1 + 30 + 600 + 10 + 200 + 80 + 10 + 40 + 1 + 6 + 100 + 70 = 1148

Finally, the value of a geminated letter (i.e., the same letter occurring twice in succession) is the same as a single letter.

---

[1] For a discussion of the sound system of Persian, see "The Sounds of Persian" in the *Tape Manual*, pp. 1-14.

[2] In the "Writing System," transliteration is employed to teach the sound-letter combinations where two or more letters represent a single sound.

[3] The Arabic numeral اوّل 'ævvæl' is also used for "first". In forming compound numbers, however, the ordinal number یکم is used more frequently. Example:

روز اوّل	ruz-e ævvæl	first day
بیست و یکم	bisto yekom	twenty-first